

SEQUENCE LISTING

<110> Vreeland, Valerie
The Regents of the University of California

<120> Recombinant Minimal Catalytic Vanadium Haloperoxidases
and Their Uses

<130> 023070-087110US

<140> US 09/596,794
<141> 2000-06-19

<150> US 09/151,189
<151> 1998-09-10

<160> 20

<170> PatentIn Ver. 2.1

<210> 1
<211> 2931
<212> DNA
<213> Fucus distichus

<220>
<221> CDS
<222> (228)..(2258)
<223> vanadium bromoperoxidase

<400> 1
cgccgacaag ccttggaga gaggttgccc aattcaacag agcgaggccc gtgaagggtgt 60
ggaggacacg tgctacaaggc tgatccacga gaggcctcaac ttccctactg atacgggagt 120
ttgtactgcg ccgcgttgcc aaaaaccgca actttaaaca gcgctcgcgaa gcgccacatg 180
cttccccacgc atccacaaaa tcgacagtgg tatcgctgag cttgaat atg ctt tgc 236
Met Leu Cys
1
cat gca gcg gac acg aca aga ggc tct cct atg cct gac acc gga gtg 284
His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp Thr Gly Val
5 10 15
ctt cgg ttg ctc aca tca gag cag cgc gct aaa ggt tgg aga cgc cag 332
Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp Arg Arg Gln
20 25 30 35
tta gag ggg gag aaa tca cta ggt ttt cat cca agc gag acg cct tat 380
Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu Thr Pro Tyr
40 45 50
atc aag tac ttg gaa ggc tct gag act tgg aag aag gtt aag ctt cca 428
Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val Lys Leu Pro
55 60 65
acg gac ggc ata tcg gct tcc aag atc ctg ggt aaa att atg gcc agg 476
Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile Met Ala Arg
70 75 80

gtc cgc atc gct acc gcc ttg gct gtg gta ctg gcc gca ccc tgt ttg Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala Pro Cys Leu	524
85 90 95	
gca ttc gac gag gtc aca gcc agt ggt gtt ttc cct gag gaa cac aag Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu Glu His Lys	572
100 105 110 115	
cac acc ggg gag gga aga cac ctc cag acc tgt aca aac tcc gac gat His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn Ser Asp Asp	620
120 125 130	
gcg ctg gat ccg acg gcg ccg aat aga agg gac aac gta gct ttt gcg Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val Ala Phe Ala	668
135 140 145	
tcg cgg cgc gat gcc gcc agg cga gaa cgt gac ggg aca ggg act gtc Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr Gly Thr Val	716
150 155 160	
tgc caa atc act aac gga gaa act gat ttg gct acc atg ttc cac aag Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met Phe His Lys	764
165 170 175	
tct ctg cca cac gat gaa ctg gga cag gta acc gca gac gac ttc gct Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp Asp Phe Ala	812
180 185 190 195	
atc ctc gag gac tgc atc tta aac gga gat ttc agc att tgc gag gac Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile Cys Glu Asp	860
200 205 210	
gtg cct gcg gga gac ccg gcg ggt cgc ctc gtc aat cct acc gct gcg Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro Thr Ala Ala	908
215 220 225	
ttt gcc atc gac ata tcc ggt ccc gca ttc tcg gct acg aca ata ccc Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr Thr Ile Pro	956
230 235 240	
ccg gta cct acc ctt tcc tct cct gag ctc gcc gct cag ttg gcg gag Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln Leu Ala Glu	1004
245 250 255	
cta tac tgg atg gcg ctg gcc agg gat gta ccc ttt atg cag tat ggc Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met Gln Tyr Gly	1052
260 265 270 275	
acc gac gaa att acc act acc gcg gca gcc aac ctc gct gga atg gga Thr Asp Glu Ile Thr Thr Ala Ala Ala Asn Leu Ala Gly Met Gly	1100
280 285 290	
ggc ttc cca aat ctg gac gcc gtg tcg ata ggg tcc gat ggt acg gtg Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp Gly Thr Val	1148
295 300 305	
gac ccg ttc tcc cag ctc ttc cga gcg acc ttc gtt ggt gtt gaa acg Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly Val Glu Thr	1196
310 315 320	

ggg ccc ttt gtc tct cag ctg ctc gtg aac agc ttc acc atc gac gct		1244	
Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr Ile Asp Ala			
325	330	335	
att acg gtc gaa ccg aag cag gag aca ttc gcc ccc gac ttg aac tat		1292	
Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp Leu Asn Tyr			
340	345	350	355
atg gtc gat ttt gac gaa tgg ctg aac att cag aat ggt gga ccc ccg		1340	
Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly Gly Pro Pro			
360	365	370	
gcc ggc ccc gaa gag tta gac gaa gag ctg cgt ttt atc cgt aac gcc		1388	
Ala Gly Pro Glu Glu Leu Asp Glu Leu Arg Phe Ile Arg Asn Ala			
375	380	385	
cgc gac ctg gcc agg gtc tcc ttc gtg gac aat atc aac acc gaa gct		1436	
Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn Thr Glu Ala			
390	395	400	
tat cgc ggg tct ctt atc cta ctt gag ctg gga gcc ttc agc agg ccc		1484	
Tyr Arg Gly Ser Leu Ile Leu Glu Leu Gly Ala Phe Ser Arg Pro			
405	410	415	
ggt atc aac ggt cca ttc atc gac agt gat cgg cag gcg ggc ttc gtc		1532	
Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala Gly Phe Val			
420	425	430	435
aac ttc ggc acg tct cac tac ttc aga ttg ata ggt gcc gcc gag ctg		1580	
Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala Ala Glu Leu			
440	445	450	
gcg cag cgt gcc tcg tgt tac caa aag tgg cag gtg cat cga ttt gca		1628	
Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His Arg Phe Ala			
455	460	465	
cgc ccc gag gct ctc ggg ggt acc ctc cac aac acc atc gcg ggg gat		1676	
Arg Pro Glu Ala Leu Gly Thr Leu His Asn Thr Ile Ala Gly Asp			
470	475	480	
cta gat gca gac ttc gac atc tcc ctt gaa aat gat gag ctc ttg		1724	
Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp Glu Leu Leu			
485	490	495	
aaa cgt gtg gcg gag ata aat gcg gcg cag aat ccc aac aac gag gtc		1772	
Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn Asn Glu Val			
500	505	510	515
acc tac ctt ctt cca caa gct atc caa gtg gga tcg cca acg cac cct		1820	
Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro Thr His Pro			
520	525	530	
tcc tac ccg tcc ggc cac gct acc caa aat gga gca ttt gcc aca gtt		1868	
Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe Ala Thr Val			
535	540	545	
ctg aag gcc ctc att ggc cta gat cggt gga ggt gag tgc ttc cct aac		1916	
Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys Phe Pro Asn			
550	555	560	

ccc gtg ttc cca agc gat gac ggc ctg gaa cta atc aac ttc gaa ggg Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn Phe Glu Gly 565 570 575	1964
gca tgc ctt aca tat gag gga gag atc aac aag ctc gcg gtc aac gtc Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala Val Asn Val 580 585 590 595	2012
gca ttt ggg agg cag atg ctg ggc atc cac tat cgg ttc gac ggt atc Ala Phe Gly Arg Gln Met Leu Gly Ile His Tyr Arg Phe Asp Gly Ile 600 605 610	2060
caa ggc cta ctt ctc gga gag aca atc act gta cga aca ctt cac cag Gln Gly Leu Leu Gly Glu Thr Ile Thr Val Arg Thr Leu His Gln 615 620 625	2108
gag ctg atg acg ttc gcc gag gaa gcc acc ttt gaa ttc cgc tta ttc Glu Leu Met Thr Phe Ala Glu Ala Thr Phe Glu Phe Arg Leu Phe 630 635 640	2156
acc gga gag gtc atc aaa ctt ttc cag gac ggg aca ttc tcc atc gat Thr Gly Glu Val Ile Lys Leu Phe Gln Asp Gly Thr Phe Ser Ile Asp 645 650 655	2204
gga gat atg tgt tcc ggt ttg gtt tac act ggc gtg gcg gac tgc cag Gly Asp Met Cys Ser Gly Leu Val Tyr Thr Gly Val Ala Asp Cys Gln 660 665 670 675	2252
gct tagtgcagaa aataataatt gtcggatgct taaaatgcac ccacgaccaa Ala	2305
gtcgtcgagt cacgtcgccg gagcatcctt cagcgaaaaa ggagagtaac ctatatgcta tagaggagaa ccacggagta caatgcaggt tctttacca tgtacattgg attgcagtaa gtgcggtag agagggatac gttaaacgtg cttgcctgtg tatatgatac atttgtcatg gaaatattag aatgcgttga cttgacttca ccatgaaata ccatgatcgc gtgggtgtgct gtttcacct gtcggagcgg tacgtaagat gtgtttcta ctgagccgtt tgtgttttagt ccattcccgcg tggcagtgt aacaaagagg atgttgtctc gccctcagtt tggagagta cgtaggtggc aggacgtata tctctggtag cggctgtta agaacttcca caagaccgtt tacgtttggc tggttagtgc atgcctcttc gttacttgac cgatccattg agagtacctg taccagtatg gtgtaagaca tattttctc ctgttatgga tctgtagaac agcttaggtgt tgtttatac acaggatgct ataaaatagg gatgttgata atggcatcg tactcatgaa accgcaaaat ggcgatagat attccc	2365 2425 2485 2545 2605 2665 2725 2785 2845 2905 2931

<210> 2
<211> 676
<212> PRT
<213> Fucus distichus

<400> 2
 Met Leu Cys His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp
 1 5 10 15

Thr Gly Val Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp
 20 25 30

Arg Arg Gln Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu
 35 40 45

Thr Pro Tyr Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val
 50 55 60

Lys Leu Pro Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile
 65 70 75 80

Met Ala Arg Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala
 85 90 95

Pro Cys Leu Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu
 100 105 110

Glu His Lys His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn
 115 120 125

Ser Asp Asp Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val
 130 135 140

Ala Phe Ala Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr
 145 150 155 160

Gly Thr Val Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met
 165 170 175

Phe His Lys Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp
 180 185 190

Asp Phe Ala Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile
 195 200 205

Cys Glu Asp Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro
 210 215 220

Thr Ala Ala Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr
 225 230 235 240

Thr Ile Pro Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln
 245 250 255

Leu Ala Glu Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met
 260 265 270

Gln Tyr Gly Thr Asp Glu Ile Thr Thr Ala Ala Ala Asn Leu Ala
 275 280 285

Gly Met Gly Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp
 290 295 300

Gly Thr Val Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly
 305 310 315 320

Val Glu Thr Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr
 325 330 335

 Ile Asp Ala Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp
 340 345 350

 Leu Asn Tyr Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly
 355 360 365

 Gly Pro Pro Ala Gly Pro Glu Glu Leu Asp Glu Glu Leu Arg Phe Ile
 370 375 380

 Arg Asn Ala Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn
 385 390 395 400

 Thr Glu Ala Tyr Arg Gly Ser Leu Ile Leu Leu Glu Leu Gly Ala Phe
 405 410 415

 Ser Arg Pro Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala
 420 425 430

 Gly Phe Val Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala
 435 440 445

 Ala Glu Leu Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His
 450 455 460

 Arg Phe Ala Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile
 465 470 475 480

 Ala Gly Asp Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp
 485 490 495

 Glu Leu Leu Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn
 500 505 510

 Asn Glu Val Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro
 515 520 525

 Thr His Pro Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe
 530 535 540

 Ala Thr Val Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys
 545 550 555 560

 Phe Pro Asn Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn
 565 570 575

 Phe Glu Gly Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala
 580 585 590

 Val Asn Val Ala Phe Gly Arg Gln Met Leu Gly Ile His Tyr Arg Phe
 595 600 605

 Asp Gly Ile Gln Gly Leu Leu Gly Glu Thr Ile Thr Val Arg Thr
 610 615 620

 Leu His Gln Glu Leu Met Thr Phe Ala Glu Glu Ala Thr Phe Glu Phe
 625 630 635 640

Arg Leu Phe Thr Gly Glu Val Ile Lys Leu Phe Gln Asp Gly Thr Phe
 645 650 655
 Ser Ile Asp Gly Asp Met Cys Ser Gly Leu Val Tyr Thr Gly Val Ala
 660 665 670
 Asp Cys Gln Ala
 675

<210> 3
 <211> 51
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:probe for
 second conserved region between Curvularia and
 Ascophyllum vanadium peroxidase active sites

<400> 3 51
 ccaacgcacc ctgcgtaccc gtctggccac gctacccaaa acggaggcatt t

<210> 4
 <211> 51
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:probe for third
 conserved region between Curvularia and
 Ascophyllum vanadium peroxidase active sites

<400> 4 51
 ccgtacgaac acttcaccag gagctgatga cttcgccga ggaatccacc t

<210> 5
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:Fucus
 peroxidase ligation independent cloning (LIC) 5'
 primer for full length construct

<400> 5 34
 gacgacgaca atatgcttg ccatgcagcg gaca

<210> 6
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:Fucus
 peroxidase ligation independent cloning (LIC) 5'
 primer for mid length construct

<pre><400> 6 gacgacgaca agatggcgcc gaatagaagg gacaa</pre> <pre><210> 7 <211> 33 <212> DNA <213> Artificial Sequence</pre> <pre><220> <223> Description of Artificial Sequence:Fucus peroxidase ligation independent cloning (LIC) 5' primer for short construct</pre> <pre><400> 7 gacgacgaca agatgctttt ccgagcgacc ttc</pre> <pre><210> 8 <211> 33 <212> DNA <213> Artificial Sequence</pre> <pre><220> <223> Description of Artificial Sequence:Fucus peroxidase ligation independent cloning (LIC) 3' primer for full length, mid length and short constructs</pre> <pre><400> 8 gaggagaagc ccgggttgcac taaggctggc agt</pre> <pre><210> 9 <211> 22 <212> PRT <213> Artificial Sequence</pre> <pre><220> <223> Description of Artificial Sequence:conserved Fucus vanadium-binding region 1, amino acids 452-473, 1st conserved motif</pre> <pre><400> 9 Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His Arg Phe Ala 1 5 10 15</pre> <pre>Arg Pro Glu Ala Leu Gly 20</pre> <pre><210> 10 <211> 19 <212> PRT <213> Artificial Sequence</pre> <pre><220> <223> Description of Artificial Sequence:conserved Fucus vanadium-binding region 2, amino acids 528-546, 2nd conserved motif</pre>	35 33 33 33 1
--	---

```

<400> 10
Pro Thr His Pro Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala
    1           5           10          15

Phe Ala Thr

<210> 11
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:conserved Fucus
      vanadium-binding region 3, amino acids 591-609,
      3rd conserved motif

<400> 11
Asn Lys Leu Ala Val Asn Val Ala Phe Gly Arg Gln Met Leu Gly Ile
    1           5           10          15

His Tyr Arg Phe Asp
    20

<210> 12
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:1st conserved
      motif from algal haloperoxidases

<220>
<221> MOD_RES
<222> (1)..(22)
<223> Xaa = any amino acid

<400> 12
Ala Xaa Xaa Xaa Xaa Xaa Tyr Gln Lys Xaa Xaa Xaa His Arg Xaa Xaa
    1           5           10          15

Arg Pro Glu Ala Xaa Gly
    20

<210> 13
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:1st conserved
      motif from algal and fungal haloperoxidases

<220>
<221> MOD_RES
<222> (2)..(11)
<223> Xaa = any amino acid

```

```

<220>
<221> MOD_RES
<222> (12)
<223> Xaa = any amino acid, present in algal enzymes only

<220>
<221> MOD_RES
<222> (13)..(19)
<223> Xaa = any amino acid

<220>
<221> MOD_RES
<222> (20)
<223> Xaa = any amino acid, present in algal enzymes only

<400> 13
Ala Xaa Xaa Xaa Xaa Xaa Xaa Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    1           5           10          15

Arg Pro Xaa Xaa Xaa Gly
    20

<210> 14
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:2nd conserved
      motif from algal haloperoxidases

<220>
<221> MOD_RES
<222> (1)..(16)
<223> Xaa = any amino acid

<400> 14
Pro Xaa His Pro Ser Tyr Xaa Ser Gly His Ala Xaa Xaa Xaa Gly Ala
    1           5           10          15

<210> 15
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:2nd conserved
      motif from algal and fungal haloperoxidases

<220>
<221> MOD_RES
<222> (1)..(16)
<223> Xaa = any amino acid

<400> 15
Pro Xaa Xaa Pro Xaa Tyr Xaa Ser Gly His Ala Xaa Xaa Xaa Gly Ala
    1           5           10          15

```

<210> 16
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:2nd conserved motif from algal and fungal haloperoxidases, phosphatases and other proteins

<220>
<221> MOD_RES
<222> (1)
<223> Ser conserved in some proteins

<220>
<221> MOD_RES
<222> (2)
<223> Tyr conserved in some proteins

<400> 16
Ser Tyr Pro Ser Gly His
1 5

<210> 17
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:3rd conserved motif from algal haloperoxidases

<220>
<221> MOD_RES
<222> (1)..(21)
<223> Xaa = any amino acid

<400> 17
Asn Lys Leu Ala Xaa Asn Xaa Ala Xaa Gly Arg Xaa Met Xaa Gly Xaa
1 5 10 15

His Tyr Xaa Xaa Asp
20

<210> 18
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:3rd conserved motif from algal and fungal haloperoxidases

<220>
<221> MOD_RES
<222> (1)..(14)
<223> Xaa = any amino acid

<400> 18
Ala Xaa Xaa Arg Xaa Xaa Xaa Gly Xaa His Xaa Xaa Xaa Asp
1 5 10

<210> 19
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:3rd conserved
motif from algal and fungal haloperoxidases,
phosphatases and other proteins

<220>
<221> MOD_RES
<222> (2)..(9)
<223> Xaa = any amino acid

<220>
<221> MOD_RES
<222> (4)
<223> Gly conserved in some proteins

<220>
<221> MOD_RES
<222> (10)
<223> Asp conserved in some proteins

<400> 19
Arg Xaa Xaa Gly Xaa His Xaa Xaa Xaa Asp
1 5 10

<210> 20
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:6xHis
purification tag

<400> 20
His His His His His His
1 5